

---

**NLSY97 QUESTIONNAIRE AND  
INTERVIEWER'S REFERENCE MANUAL  
ROUND 2**

---

**Prepared for the  
U.S. Department of Labor by**

**Center for Human Resource Research  
The Ohio State University  
2000**



## TABLE OF CONTENTS

---

<b>Introduction .....</b>	<b>i</b>
I. Types and Formats of Question Blocks.....	ii
II. Instrument Compression .....	v
III. Save Array.....	vi
 <b>Round 2 Questionnaire.....</b>	 <b>1</b>
Household Information .....	3
Schooling .....	31
Employment .....	99
Training.....	191
Time Use.....	207
Health .....	209
Self-Administered .....	210
Marriage .....	280
Fertility .....	296
Program Participation.....	309
Income .....	396
Assets.....	413
PIAT Mathematics Assessment .....	490
Locator.....	491
Interviewer Remarks .....	497
 <b>Household Income Update.....</b>	 <b>503</b>
 <b>Interviewer Reference Manual .....</b>	 <b>507</b>
Introduction.....	509
Household Information .....	509
Schooling .....	510
Employment.....	512
Training .....	523
Health .....	525
Marriage .....	525
Fertility .....	526
Program Participation.....	527
Income .....	528
Assets.....	529

This publication is prepared in conjunction with contract # J-9-J-9-0007 with the Bureau of Labor Statistics, U.S. Department of Labor. None of its contents is to be construed as necessarily representing the official position or policy of the Department of Labor.

For more information about any aspect of the NLS program, contact:

NLS User Services  
Center for Human Resource Research  
921 Chatham Lane, Suite 100  
Columbus, OH 43221-2418  
(614) 442-7366  
usersvc@postoffice.chrr.ohio-state.edu



---

## INTRODUCTION

This document describes the content of the Round 2 interview of the NLSY97 cohort, conducted in October 1998–May 1999. This survey used two different questionnaires: the Youth Questionnaire and the Household Income Update. The Youth Questionnaire was administered to the NLSY97 respondent, and the one-page paper Household Income Update was filled out by a parent of non-independent youth respondents. Both instruments are included in this document.

Except for the Household Income Update, this survey was conducted entirely as a computer-assisted interview. Although some cases were done over the telephone, the dominant interview mode was the in-person interview, so we will use the term CAPI (computer-assisted personal interview) to refer to the data collection technology.

Our goal is to make the content of the interview as comprehensible as possible by producing a clear, complete, and compact questionnaire. With this in mind, the Round 2 CAPI questionnaires have been formatted as similarly as possible to a conventional questionnaire. They also include some additional information that enhances the usefulness of each questionnaire for data users. The resulting documents accurately summarize the “conversation” between the interviewer and respondent. Suggestions for improving the questionnaire or comments on errors or inconsistencies in the documentation should be directed to CHRR User Services.

We use the term “conversation” to emphasize the difference between the questionnaire and the codebook. The codebook contains the data released as a result of the interview, but some information collected during the interview is not present in the codebook. For example, the questionnaire includes questions that collected names of household members, employers, etc. Such items are not released, so the questions that collect them are not present in the codebook. Conversely, the codebook contains some information not directly collected in the interview. For example, while created variables are documented in the codebook, there is no reference to them in the questionnaire as they are not part of the conversation between the interviewer and respondent.

To aid researchers in using the questionnaire, this introduction provides information on several topics. First, typical question blocks are described and key terms such as “R Number,” “Distribution Code Block,” and “Lead in Questions” are defined. This section explains how to follow the flow of the interview through the questionnaire. Second, the instrument compression techniques that have helped us to produce a compact questionnaire are described. Finally, this introductory section concludes with a discussion of save arrays, which provide interviewers with information specific to a given respondent (e.g., employer, name).

## I. TYPES AND FORMATS OF QUESTION BLOCKS

Question records are the building blocks of the CAPI questionnaires, with various types of records performing different data collection and/or manipulation functions. Loosely grouped, CAPI question records will either: (a) provide administrative information that assists the interviewer or provides an introduction to the next question, (b) provide question text and control for data entry of the response, or (c) perform internal operations and calculations to guide the interview and maintain the internal data storage during the interview. All question records essential to understanding either the content of the survey or the flow of respondents through the interview will appear within these Round 2 NLSY97 questionnaires.

To aid users, the Round 2 questionnaire contains some elements found in the Round 2 NLSY97 codebook. For instance, codebook reference numbers are provided for question items and additional information such as universe skip patterns are included in each question block. Additionally, many of the elements not essential to understanding the conversational flow of the interview have been removed, making each question block less cluttered and easier to read.

Figure 1 shows the format of a typical question block; the various information fields are explained in the following paragraphs.

Question name [R number]	
Question text	Question Text?
Question text	(INTERVIEWER COMMENTS)
Question text	internal code;
Question text	/* comments */
Question text	(space for continuation)
(Code text of Distribution Code Block                      Skip)	
K1 distribution code block text1	(GO TO YSCH-nn)
K2 distribution code block text2	
K2 distribution code block text3	
.	
.	
.	
Kn distribution code block textn	
Lead In(s):	
Default Next Question:	

**Figure 1. CAPI Question Block**

### A. Question Name

Question names are designed to be the unique identifiers assigned to identical questions across survey years. Each question in the Youth Questionnaire has a four letter base before the number (e.g., YSCH-1600). The first letter, Y, indicates that this is a question asked of the youth respondents. This questionnaire is subdivided into the following sections: Household Information (YHHI), Schooling (YSCH), Employment (YEMP), Training (YTRN), Time Use (YTIM), Health (YHEA), Self-Administered (YSAQ), Marriage

(YMAR), Fertility (YFER), Program Participation (YPRG), Income, (YINC), Assets (YAST), PIAT Math (YPIA), Locator (YLOC), and Interviewer Remarks (YIR).

The question names in the separate Household Income Update begin with HIU.

## B. R Number

The NLSY97 identifies the variables by reference numbers (R numbers). The identifier for a particular variable in a particular year does not change, even as new waves are released. The R number has the format Rxxxxx. Every question in the questionnaire which appears in the codebook has its reference number printed next to it as an identification. This allows the user to identify immediately which items he or she will be able to extract from the CD and to find a given question on the CD with a minimum amount of effort. Any item that is not in the public use data set will not have an R number in this questionnaire; these items are identified by question name only.

Most questions generate a single variable in the codebook; some, however, will generate more than one variable. For example, some questions collect information about the date an event happened and split this information into three separate variables: month, day and year. Some questions offer a respondent several answers to choose from with the instruction to choose all that apply, generating a great many variables. In these questions, each possible outcome is a distinct variable, coded as “selected,” “not selected,” or “valid skip.”

## C. Question Text

This field contains the text of the question that the interviewer asked the respondent or other information used to direct the flow of the interview. The following types of text may appear:

Question Text:	Text that is asked of the respondent appears in a conventional mixture of upper and lowercase letters.
INTERVIEWER COMMENTS:	Text in all uppercase is either an instruction to the interviewer or a clarifying comment. NORC interviewers do not read these items to the respondent.
Internal Code:	When a question record contains a machine instruction, that instruction is printed in this field. Such instructions end with a semicolon. These instructions are accompanied by internal comments as described below.
/* comments */	Internal comments that explain what a machine instruction does are set off by a leading “/*” and a trailing “*/.” These comments explain the function being performed by the internal code so that persons unfamiliar with the command system can follow the interview protocol.

## D. Distribution Code Block

When a question requires the choice of one or more items from a predefined list, that list, or distribution code block, is shown in the question block. Each item in the distribution code block has three components:

Code:	The numeric code associated with each possible response is listed first. All data in the NLSY97 public use data file are numeric, so each possible item that could be selected is associated with an integer.
Text:	Next is the item text or description of the code, such as “Yes” or “No.” If the question uses a “show card” (this will be indicated in the question text field), the items on the show card are sometimes identified by letters or numbers; these letters or numbers are included with the response text in this questionnaire.
Skip:	When the selection of a specific response determines which question will be asked next, a ( <i>GO TO Rxxxxx</i> ) or ( <i>GO TO YSCH-nn</i> ) notation follows the response text.

For some questions that collect numeric data (hours, weeks, dollar amounts, etc.), the question block contains what appears to be a distribution code block. This is present whenever the next question to be asked depends upon the numeric value falling in some range. In such cases the distribution code block is constructed so it corresponds to these skip conditions. However, the interviewer would actually enter the exact number provided by the respondent and this exact number would be recorded in the data set.

## E. Lead in Questions

The CAPI questionnaires include a “Lead in(s)” field in the question block. The field lists the questions that lead into each particular item. By tracing the skip pattern backward, the user is able to determine the universe of respondents asked a given question. Examples:

YSCH-1200 [Default]	This means that the default path from question YSCH-1200 leads to the current question, but there may be conditions under which a different path would be taken.
YSCH-1200 [1:4]	This means that whenever the response category for question YSCH-1200 takes on the values one to four (inclusive), the next question will be the current question record. If the response is some other value, the respondent may next be asked a different question.

## F. Default Next Question

This field specifies the question to be asked next unless another skip specification in the distribution code block applies. If the default next question is the following question in the printed questionnaire, no skip specification appears. Otherwise, a “Go to” message at the bottom of the question record specifies the default next question.



## II. INSTRUMENT COMPRESSION

---

In order to provide a more compact questionnaire, two types of records have been dropped. The following do not appear within this questionnaire: (1) “looped” questions or repetitive series and (2) question records that perform internal operations not necessary for understanding the conversational flow of the survey.

First, repetitions of questions that asked multiple times are not included. Questions appearing in loops will be enclosed inside a special type of question called a REPEAT command, which begins a loop, and the UNTIL command, which will end the loop. For example, the CAPI instrument asks questions about the education, employment, etc., for up to twenty persons in the household. Instead of printing all twenty sets of questions, only the questions for the first person are printed.

Second, the instrument was compressed by eliminating question records that contained only nonessential skip pattern information. The Round 2 Youth CAPI instrument included 5,552 numbered question records, only some of which are essential to understanding the action content and conversational flow of the survey. The process of dropping unessential records and then changing the lead in and next question specifications so that the skip pattern does not have any gaps is referred to as “skip pattern compression.” The resulting questionnaire still contains all relevant information about skip patterns and universes and is easier for researchers to use.

All questions that were part of the original questionnaire and appear as variables on the Round 2 CD have been included in this questionnaire.

### III. SAVE ARRAY

---

“Save arrays” are fields in which information is stored. Each such field, or save array, is assigned a name. For example, the name of the respondent’s current employer would be stored in the “EMPLOYER.NAME(1)” save array for use during the interview. This “save array name” is then used to reference and access the data in the save array field during the actual operation of the survey. Each time a question refers to the current employer, the computer automatically inserts the name of the employer into the question. Information stored in save arrays is used:

- a) as part of the actual text of survey questions (e.g., the name of a given household member, child, employer, types of training programs, different sources of income for household members);
- b) to govern paths through the questionnaire. The save array field is accessed, and the content (or absence of content) is checked. The path through the questionnaire is determined based upon the information in the save array field; and
- c) as elements in equations and calculations. These produce new pieces of information that may be incorporated later in the instrument.

When used in a question text, save arrays are similar to the text fills used in PAPI interviews (a paper and pencil interviewing method used in previous NLS age cohort studies). These text fills indicated that the interviewer should substitute a given word or phrase into the question at the time of the interview so it read correctly. For example, in a printed PAPI questionnaire one would often see “*Since (DATE OF LAST INTERVIEW) have you...*”; this phrase would instruct the interviewer to insert the actual date for a given respondent. However, correctly inserting such text fills was the job of the interviewer and often involved flipping between pages.

One immediate difference in the CAPI questionnaires is that there are far more text fills, and most of them are automated or handled by the computer. Some CAPI text fills are identical to those used in the PAPI interviews, such as “*he/she*” to handle gender fills. For date-related text fills, the system substitutes a previously supplied date for the descriptive text specifying which date is needed.

In this questionnaire, text fills that refer to a save array are represented by bracketed text describing what piece of information is substituted in the interview. Although the interviewer never sees this bracketed text, researchers can use that information to follow the flow of the interview.